Drug target validation and identification of secondary drug target effects using DNA microarrays

MATTHEW J. MARTON[1], JOSEPH L. DERISI[2], HOLLY A. BENNETT[1], VISHWANATH R. IYER[2], MICHAEL R. MEYER[1], CHRISTOPHER J. ROBERTS[1], ROLAND STOUGHTON[1], JULJA BURCHARD[1], DAVID SLADE[1], HONGYUE DAI[1], DOUGLAS E. BASSETT, JR[1]., LELAND H. HARTWELL[3], PATRICK O. BROWN[2] & STEPHEN H. FRIEND[1]

[1]Rosetta Inpharmatics, 12040 115th Avenue NE, Kirkland, Washington 98034, USA
[2]Department of Biochemistry, Stanford University School of Medicine, Howard Hughes Medical Institute Stanford, California 94305-5428, USA
[3]Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue N., Seattle, Washington 98109, USA
Correspondence should be addressed to S.H.F; email: sfriend@rosetta.org

We describe here a method for drug target validation and identification of secondary drug target effects based on genome-wide gene expression patterns. The method is demonstrated by several experiments, including treatment of yeast mutant strains defective in calcineurin, immunophilins or other genes with the immunosuppressants cyclosporin A or FK506. Presence or absence of the characteristic drug 'signature' pattern of altered gene expression in drug-treated cells with a mutation in the gene encoding a putative target established whether that target was required to generate the drug signature. Drug dependent effects were seen in 'targetless' cells, showing that FK506 affects additional pathways independent of calcineurin and the immunophilins. The described method permits the direct confirmation of drug targets and recognition of drug-dependent changes in gene expression that are modulated through pathways distinct from the drug's intended target. Such a method may prove useful in improving the efficiency of drug development programs.

Good drugs are potent and specific; that is, they must have strong effects on a specific biological pathway and minimal effects on all other pathways. Confirmation that a compound inhibits the intended target (drug target validation) and the identification of undesirable secondary effects are among the main challenges in developing new drugs. Comprehensive methods that enable researchers to determine which genes or activities are affected by a given drug might improve the efficiency of the drug discovery process by quickly identifying potential protein targets, or by accelerating the identification of compounds likely to be toxic. DNA microarray technology, which permits simultaneous measurement of the expression levels of thousands of genes, provides a comprehensive framework to determine how a compound affects cellular metabolism and regulation on a genomic scale[1-11]. DNA microarrays that contain essentially every open reading frame (ORF) in the *Saccharomyces cerevisiae* genome have already been used successfully to explore the changes in gene expression that accompany large changes in cellular metabolism or cell cycle progression[7-10].

In the modern drug discovery paradigm, which typically begins with the selection of a single molecular target, the ideal inhibitory drug is one that inhibits a single gene product so completely and so specifically that it is as if the gene product were absent. Treating cells with such a drug should induce changes in gene expression very similar to those resulting from deleting the gene encoding the drug's target. Here we have compared the genome-wide effects on gene expression that result from deletions of various genes in the budding yeast *S. cerevisiae* to the effects on gene expression that result from treatment
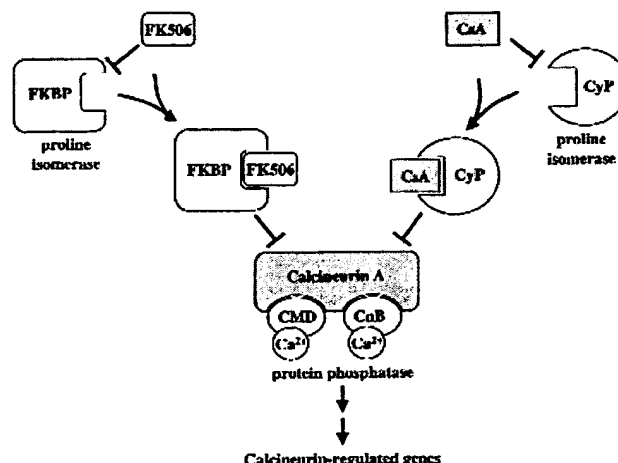
with known inhibitors of those gene products. Using the calcineurin signaling pathway as a model system, we tested an approach that permits identification of genes that encode proteins specifically involved in pathways affected by a drug. The FK506 characteristic pattern, or 'signature', of altered gene expression was not observed in mutant cells lacking proteins inhibited by FK506 (for example, a calcineurin or FK506-binding-protein mutant strain), but was observed in mutants deleted for genes in pathways unrelated to FK506 action (for example, a cyclophilin mutant strain). Conversely, the cyclosporin A (CsA) signature was not observed in CsA-treated calcineurin or cyclophilin mutant strains, but was seen in an FK506-binding-protein mutant strain treated with CsA. The method also demonstrates that FK506, a clinically used immunosuppressant, has 'off-target' effects that are independent of its binding to immunophilins. Thus, the approach we describe may provide a way to identify the pathways altered by a drug and to detect drug effects mediated through unintended targets.

Null mutants phenocopy drug-treated cells on a genomic scale

To test whether a null mutation in a drug target serves as a model of an ideal inhibitory drug, we examined the effects on gene expression associated with pharmacological or genetic inhibition of calcineurin function. Calcineurin is a highly conserved calcium- and calmodulin-activated serine/threonine protein phosphatase implicated in diverse processes dependent on calcium signaling[12-13]. In budding yeast, calcineurin is required for intracellular ion homeostasis[14], for adaptation to prolonged mating pheromone treatment[15] and in the regulation of

ARTICLES ⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯

Fig. 1 Model of antagonism of the calcineurin signaling pathway mediated by FK506 and cyclosporin A (CsA). Calcineurin activity is composed of a catalytic subunit (calcineurin A, encoded in yeast by the *CNA1* and *CNA2* genes), and calcium-binding regulatory subunits calmodulin (CMD) and calcineurin B (CnB). After entering cells, FK506 and CsA specifically bind and inhibit the peptidyl-proline isomerase activity of their respective immunophilins, FK506 binding proteins (FKBP) and cyclophilins (CyP). The most abundant immunophilins in yeast (Fpr1 and Cph1) are thought to mediate calcineurin inhibition. Drug–immunophilin complexes bind and inhibit the calcium- and calmodulin-stimulated phosphatase calcineurin. Among the substrates of calcineurin are transcriptional activators that act to modulate gene expression.

the onset of mitosis[16]. In mammals, calcineurin has been implicated in T-cell activation[12], in apoptosis[17], in cardiac hypertrophy[18] and in the transition from short-term to long-term memory[19]. In both organisms, calcineurin activity is inhibited by FK506 and CsA, immunosuppressant drugs whose effects on calcineurin are mediated through families of intracellular receptor proteins called immunophilins[12,20] (Fig. 1). To assess the effects of pharmacologic inhibition of calcineurin, wild-type *S. cerevisiae* was grown to early logarithmic phase in the presence or absence of FK506 or CsA. Isogenic cells, from which the genes encoding the catalytic subunits of calcineurin (*CNA1* and *CNA2*) had been deleted[21] (referred to as the *cna* or calcineurin mutant), were grown in parallel, in the absence of the drug. Fluorescently-labeled cDNA was prepared by reverse transcription of polyA+ RNA in the presence of Cy3- or Cy5-deoxynucleotide triphosphates and then hybridized to a microarray containing more than 6,000 DNA probes representing 97% of the known or predicted ORFs in the yeast genome. Simultaneous hybridization of Cy5-labeled cDNA from mock-treated cells and Cy3-labeled cDNA from cells treated with 1 µg/ml FK506 allowed the effect of drug treatment on mRNA levels of each ORF to be determined (Fig. 2a and b and data not shown). Similarly, effects of the calcineurin mutations on the mRNA levels of each gene were assessed by simultaneous hybridization of Cy5-labeled cDNA from wild-type cells and Cy3-labeled cDNA from the calcineurin mutant strain(Fig. 2c). For each comparison of this kind, reported expression ratios are the average of at least two hybridizations in which the Cy3 and Cy5 fluors were reversed to remove biases that may be introduced by gene-specific differences in incorporation of the two fluors (data not shown).

Treatment with FK506 in these growth conditions resulted in a signature pattern of altered gene expression in which mRNA levels of 36 ORFs changed by more than twofold (http://www.rosetta.org). A very similar pattern of altered gene expression was observed when the calcineurin mutant strain was compared to wild-type cells. Comparison of the changes in mRNA expression of each gene resulting from treatment of wild-type cells with FK506 with mRNA expression changes resulting from deletion of the calcineurin genes showed the considerable similarity of the global transcript alterations in response to the two perturbations (Fig. 2b–d). Quantification of this similarity using the correlation coefficient (ρ) showed large correlations between the FK506 treatment signature and the calcineurin deletion signature (ρ = 0.75 ± 0.03), as well as the CsA treatment signature (ρ = 0.94±0.02), but not with a randomly selected deletion mutant strain (deleted for the *YER071C* gene; ρ = −0.07 ± 0.04; Fig. 2e). The FK506 treatment signature was also compared with those of more than 40 other deletion mutant strains or drug-treatments thought to affect
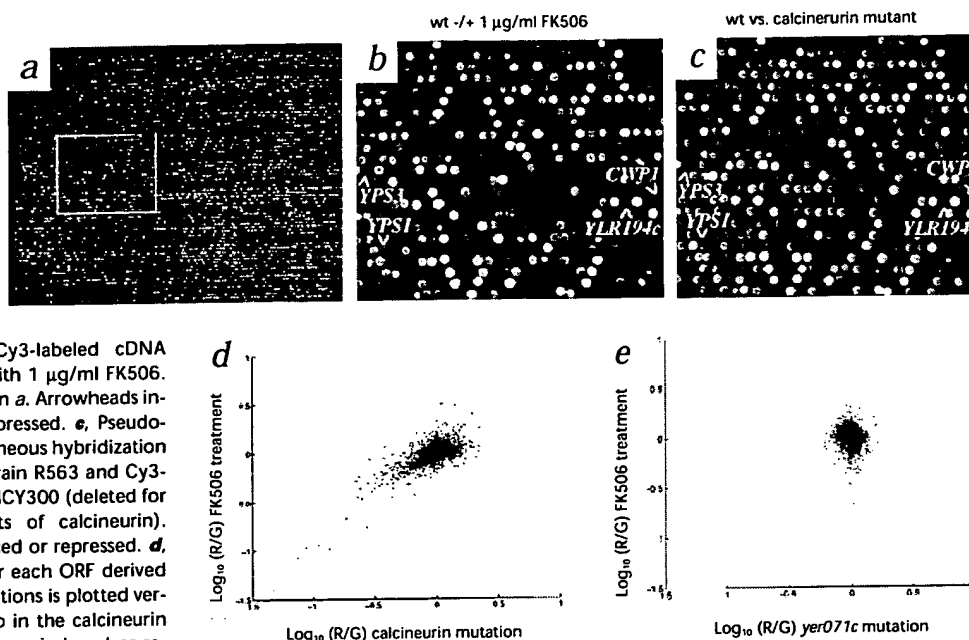
unrelated pathways, and none had statistically significant correlations. These data establish that genetic disruption of calcineurin function provides a close and specific phenocopy of treatment with FK506 or CsA.

To avoid generalizing from a single example, we also compared the effects of treatment of wild-type cells with 3-aminotriazole (3-AT) with the effects of deletion of the *HIS3* gene. *HIS3* encodes imidazoleglycerol phosphate dehydratase, which catalyzes the seventh step of the histidine biosynthetic pathway in yeast[22]; 3-AT is a competitive inhibitor of this enzyme that triggers a large transcriptional amino-acid starvation response[23]. Microarray analysis of wild-type and isogenic *his3*-deficient strains demonstrated the expected large genome-wide transcriptional responses (involving more than 1,000 ORFs) resulting from treatment with 3-AT (Fig. 3a) or from *HIS3* deletion (Fig. 3c). Quantitative comparison of the 3-AT treatment signature and the *his3* mutant signature showed a high level of correlation (ρ= 0.76 ± 0.02) that even extended to genes that experienced small changes in expression level (Fig. 3b). As a negative control, the correlations between the 3-AT treatment signature or the *his3* mutant signature and the calcineurin mutant strain were not statistically significant (ρ = 0.09 ± 0.06 and −0.01 ± 0.04, respectively). That both the calcineurin/FK506 and the *his3*/3-AT comparisons were highly correlated indicates that in many cases the expression profile resulting from a gene deletion closely resembles the expression profile of wild-type cells treated with an inhibitor of that gene's product.

'Decoder' strategy: Drug target validation with deletion mutants
Because pharmacological inhibition of different targets might give similar or identical expression profiles, simple comparison of drug signatures to mutant signatures is unlikely to unambiguously identify a drug's target. To overcome this limitation, an additional 'decoder' step is used. We first compare the expression profile of wild-type drug-treated cells to the expression profiles from a panel of genetic mutant strains, using a correlation coefficient metric. Mutant strains whose expression profile is similar to that of drug-treated wild-type cells are selected and subjected to drug treatment, generating the drug signature in the mutant strain (that is, the mutant drug signature). If the mutated gene encodes a protein involved in a pathway affected by the drug, we expect the drug signature in mutant cells to be different (or absent, for an ideal drug) from the drug signature seen in wild-type cells.

Fig. 2 Expression profiles from FK506-treated wild-type (wt) cells and a calcineurin-disruption mutant strain share a genome-wide correlation. DNA microarray analysis showing changes in gene expression resulting from FK506 treatment (*a* and *b*) or from genetic disruption of genes encoding calcineurin (*c*). *a*, Pseudo-color image of the results of simultaneous hybridization of Cy5-labeled cDNA (red) from mock-treated strain R563 and Cy3-labeled cDNA (green) from strain R563 treated with 1 μg/ml FK506. *b*, Enlarged view of the boxed area in *a*. Arrowheads indicate specific ORFs induced or repressed. *c*, Pseudo-color image of the results of simultaneous hybridization of Cy5-labeled cDNA (red) from strain R563 and Cy3-labeled cDNA (green) from strain MCY300 (deleted for the *CNA1,CNA2* catalytic subunits of calcineurin). Arrows indicate specific ORFs induced or repressed. *d*, The log$_{10}$ of the expression ratio for each ORF derived from the FK506 treatment hybridizations is plotted versus the log$_{10}$ of the expression ratio in the calcineurin mutant hybridizations. ORFs that were induced or repressed in both experiments are shown as green and red dots, respectively. *e*, The log$_{10}$ of the expression ratio for each ORF derived from the FK506 treatment hybridizations is plotted versus the log$_{10}$

of the expression ratio in the *yer071c* mutant hybridizations. No ORFs were induced or repressed in both experiments.

To illustrate this, we treated the *his3* mutant strain with 3-AT. The signature pattern of altered gene expression resulting from treatment of the mutant strain with 3-AT was much less complex than that of the 3-AT signature in wild-type cells (Fig. 4). This is seen simply by examining plots of mean intensity of the hybridization signal (which approximately reflects level of expression) versus the expression ratio for each ORF (Fig. 4). Genes that were expressed at higher or lower levels in 3-AT treated cells or in *his3* mutant cells are shown as red and green dots, respectively. We analyzed the 3-AT signature in wild-type (Fig. 4a) and *his3* mutant cells (Fig. 4c), as well as the *his3* mutant strain signature (Fig. 4b). Whereas histidine limitation induced by 3-AT induced more than 1,000 transcription-level changes in the wild-type strain, few or no transcript level changes were induced by treatment of the *his3*-deletion strain with 3-AT. This indicates that with the growth conditions used, essentially all of the effects of 3-AT depend on or are mediated through the HIS3 gene product.

Applying this approach to the calcineurin signaling pathway showed the specificity of the method. The calcineurin mutant strain and strains with deletions in the genes encoding the most abundant immunophilins in yeast[12] (*CPH1* and *FPR1*) were treated with either FK506 or CsA to determine the profiles

of altered gene expression resulting from drug treatment of the mutant cells (that is, mutant +/- drug). We compared the drug signatures in the mutants to the wild-type drug signature using the correlation coefficient metric (Table 1). Although the signature generated by treatment of wild-type cells with FK506 was highly correlated to the calcineurin mutant strain signature (ρ = 0.75 ± 0.03), it bore no similarity to the profile after treatment of the calcineurin mutant strain with FK506 (ρ = -0.01 ± 0.07). This indicates that FK506 was unable to elicit its normal transcriptional response in the calcineurin mutant strain. Likewise, treatment of the *fpr1* mutant strain with FK506 elicited an expression profile that was not correlated to the FK506 signature in the wild-type strain (ρ = -0.23 ± 0.07), indicating that the *FPR1* gene product is likely to be involved in the pathway affected by FK506. The same was true for the *cna fpr1* mutant strain. In contrast, treatment of the *cph1* mutant strain with FK506 generated an expression profile highly correlated with the wild-type FK506 expression profile (ρ = 0.79 ± 0.03), indicating the *cph1* mutation did not block the mode of action of FK506 and thus is not directly involved in the pathway affected by FK506. We tabulated the change in expression in response to FK506 in different mutant strains for all ORFs with expression ratios greater than 1.8 in FK506-treated cells or in the calcineurin mutant strain (Fig. 5a).The calcineurin mutant strain signature and the FK506 responses in wild-type and the *cph1* mutant strain are similar, and there are no transcript-level changes (seen in black) for treatment of the calcineurin, *fpr1* and *cna fpr1* mutant strains with FK506 (Fig. 5a).

Similar experiments and analyses with CsA provided further validation of this approach. The expression profile elicited by treatment of wild-type cells with CsA was highly corre-
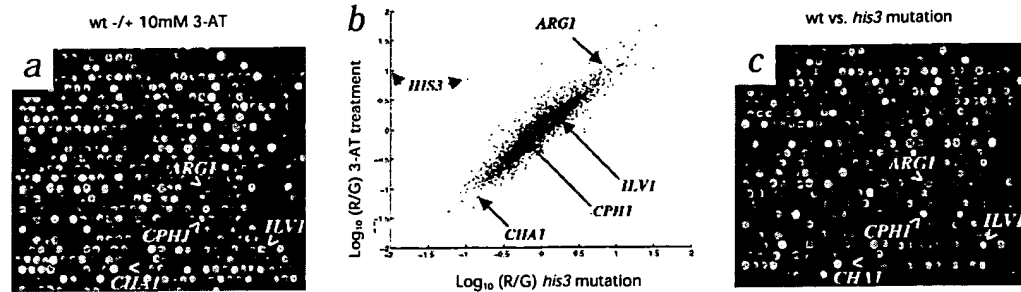
Table 1 Signature correlation of expression ratios as a result of FK506 treatment in various mutant strains

	wild-type +/-FK506	cna +/-FK506	fpr1 +/-FK506	cna fpr1 +/-FK506	cph1 +/-FK506
wild-type +/- FK506	0.93 ± 0.04	-0.01 ± 0.07	-0.23 ± 0.07	0.12 ± 0.07	0.79 ± 0.03

Signature correlation shows the absence of the FK506 signature specifically in the calcineurin (*cna*) and *fpr1* (major FK506 binding protein) deletion mutants. *cna* represents the mutant with deletions of the catalytic subunits of calcineurin, *CNA1* and *CNA2*. The correlation coefficient reported in the first column represents the correlation between two pairs of hybridizations from independent wild-type +/- FK506 experiments.

ARTICLES

Fig. 3 Expression profiles from a *his3* mutant strain and wild-type (wt) cells treated with 3-AT share a genome-wide correlation. DNA microarray analysis showing changes in gene expression resulting from 3-AT treatment (*a*) or from genetic disruption of the *HIS3* gene (*c*). *a*, Pseudo-color image of the results of simultaneous hybridization of Cy5-labeled cDNA (red) from mock-treated wild-type strain R491 and Cy3-labeled cDNA (green) from strain R491 treated with 10 mM 3-AT. *b*, Plot of the log₁₀ of the expression ratio for each ORF derived from the 3-AT treatment hybridizations is plotted versus the log₁₀ of the expression ratio in the *his3* mutant hybridizations. ORFs that were induced or repressed in both experiments are shown as green and red dots, respectively. The correlation of expression ratios applies not only to genes with large expression ratios (for example, *CHA1* and *ARG1*), but also extends to genes with expression ratios less than 2 (for example, *ILV1* and *CPH1*). *ILV1* is induced 1.9-fold and 1.5-fold, and *CPH1* is downregulated 1.9-fold

and 1.7-fold, in cells treated with 3-AT and *his3* mutant cells, respectively. Two ORFs do not fall on the line x = y. The leftmost point is the *HIS3* data point, which is induced by 3-AT treatment but which is not absent from the *his3* mutant strain. The other point is *YOR203w*. Both data points are labeled *HIS3* because hybridization to *YOR203w* is most likely due to *HIS3* mRNA, as *YOR203w* overlaps the *HIS3* open reading frame. *c*, Pseudo-color image of the results of simultaneous hybridization of Cy5-labeled cDNA (red) from wild-type strain R491 and Cy3-labeled cDNA (green) from strain R1226, deleted for the *HIS3* gene. Arrowheads indicate specific ORFs induced or repressed.

---

lated to the profile elicited by mutation of the calcineurin genes (ρ = 0.71 ± 0.04), but did not correlate with the expression profile resulting from treatment of the calcineurin mutant strain with CsA (ρ = -0.05 ± 0.07; Table 2), indicating that the genetic deletion of calcineurin interfered with the ability of CsA to elicit its normal transcriptional response. Likewise, the CsA signature was essentially absent in CsA-treated *cph1* mutant cells, and the expression profile of CsA-treated *cph1* mutant cells correlated poorly to that of CsA-treated wild-type cells (ρ = 0.18 ± 0.07). Thus, the *CPH1* gene product was required for the CsA response seen in wild-type cells. Conversely, treatment of *fpr1* mutant cells with CsA resulted in an expression pattern very similar to the profile of CsA-treated wild-type cells (ρ = 0.77 ± 0.03), indicating that *FPR1* was not necessary for the CsA-mediated effects. Analysis of individual ORFs affected by CsA and their expression ratios over the entire set of experiments confirmed that *CPH1* and the genes encoding calcineurin, but not

FPR1, are necessary for the wild-type CsA response (Fig. 5*b*). The observation that the profiles resulting from FK506 or CsA drug treatment are similar to that of the calcineurin deletion mutant strain might allow the prediction that calcineurin was involved in the pathway affected by these drugs. But because the expression profile of the *fpr1* mutant strain did not bear a strong similarity to the wild-type drug expression profile for FK506, it is obvious that the drug treatment of the mutant strains was necessary to identify Fpr1, but not Cph1, as a potential FK506 drug target. In the same way, the 'decoder' strategy was necessary to identify Cph1, but not Fpr1, as a potential drug target for CsA.

'Decoder' approach can identify secondary drug effects

For a drug that has a single biochemical target, the strategy outlined above may be useful in target validation. In many cases, however, a compound may affect multiple pathways and elicit a very complex signature. 'Decoding' such a complex signature
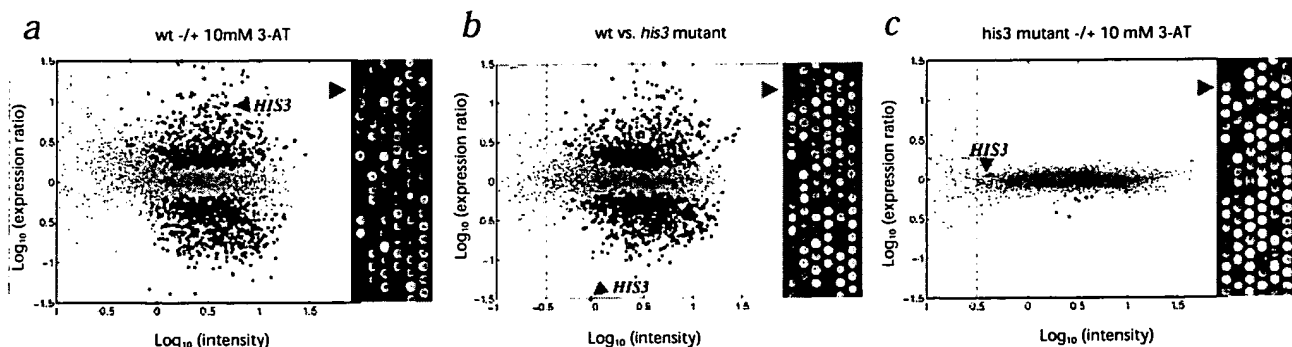
Fig. 4 Treatment of the *his3* mutant strain with 3-AT shows nearly complete loss of 3-AT signature. A plot of the log₁₀ of the mean intensity of hybridization for each ORF versus the log₁₀ of its expression ratio for each experiment is shown next to a pseudo-color image of a representative portion of the microarray. ORFs that are induced or repressed at the 95% confidence level are shown in green and red, respectively. *a*, Expression profile from treatment of the wild-type (wt) strain with 3-AT. Cy5-labeled cDNA (red) from mock-treated strain R491 and Cy3-labeled cDNA (green) from strain R491 treated with 10 mM 3-AT. *b*, Expression profile

from the *his3* deletion strain. Cy5-labeled cDNA (red) from strain R491 and Cy3-labeled cDNA (green) from strain R1226, deleted for the *HIS3* gene. *c*, Expression profile of treatment of the *his3* deletion strain with 3-AT. Cy3-labeled cDNA (red) from *his3*-deleted strain R1226 and Cy5-labeled cDNA (green) from strain R1226 treated with 10 mM 3-AT. Arrowheads indicate the DNA probe and data point corresponding to the *HIS3* gene. The blue dashed line represents the threshold below which errors tend to increase rapidly because spot intensities are not sufficiently above background intensity.

Table 2 Signature correlation of expression ratios as a result of CsA treatment in various mutant strains

	wild-type +/−CsA	cna +/−CsA	fpr1 +/−CsA	cna cph1 +/−CsA	cph1 +/−CsA
wild-type +/− CsA	0.94 ± 0.04	−0.05 ± .07	0.77 ± 0.03	−0.11 ± 0.07	0.18 ± 0.07

Signature correlation shows the absence of the CsA signature specifically in the calcineurin (*cna*) and *cph1* (cyclophilin) deletion mutants. *cna* represents the mutant with deletions of the catalytic subunits of calcineurin, *CNA1* and *CNA2*. The correlation coefficient reported in the first column represents the correlation between two pairs of hybridizations from independent wild-type +/− CsA experiments.

into the effects mediated through the intended target (the 'on-target signature') and those mediated through unintended targets (the 'off-target' signature) might be useful in evaluating a compound's specificity. Our 'decoder' strategy is based on the premise that 'off-target' signature should be insensitive to the genetic disruption of the primary target.

To determine whether the 'decoder' approach could identify an 'off-target' profile, we looked for a drug-responsive gene whose expression is insensitive to deletion of the primary target. To increase the likelihood of observing such genes, the same strains described in Tables 1 and 2 were treated with higher concentrations (50 µg/ml) of FK506. This led to a much more complex expression profile in wild-type cells, indicating that at this higher concentration, FK506 was inhibiting or activating additional targets. Several of the ORFs in this expanded FK506-induced expression profile were not affected by the calcineurin, *cph1* or *fpr1* mutations, as drug treatment of these mutant strains did not block their presence in the FK506 expression signature (Fig. 6). This indicates that FK506 was triggering changes in transcript levels of many genes through pathways independent of calcineurin, *CPH1* and *FPR1*. Many of the upregulated ORFs in the 'off-target' pathway were genes reported to be regulated by the transcriptional activator Gcn4 (ref. 24). In some strains, a reporter gene under *GCN4* control was induced in response to FK506 treatment[25]. To determine whether *GCN4* is involved in this pathway that is independent of calcineurin, *CPH1* and *FPR1*, we analyzed the effects of treatment with high-dose FK506 on global gene expression in a strain with a *GCN4* deletion (Fig. 6). Of the 41 ORFs with calcineurin-independent expression ratios greater than 4, 32 were not induced in the *gcn4* mutant, indicating that their induction by FK506 was *GCN4*-dependent. Not all *GCN4*-regulated genes were induced by FK506. This FK506-induced subset of *GCN4*-regulated genes may be those most sensitive to subtle changes in Gcn4 levels, or perhaps other regulatory circuits prevent FK506 activation of some *GCN4*-regulated genes. Seven of the remaining nine ORFs induced by FK506 were independent of

both the calcineurin and *GCN4* pathways. The simplest explanation is that FK506 inhibits or activates additional pathways. Members of this class include *SNQ2* and *PDR5*, genes that encode drug efflux pumps with structural homology to mammalian multiple drug resistance proteins[26]. FK506 may interact directly with Pdr5 to inhibit its function[27]. Our results indicate that treatment with FK506 leads to fourfold-to-sixfold induction of *PDR5* mRNA levels. *YOR1*, another gene that can confer drug resistance, is also induced threefold-to-fourfold by FK506. Thus, drug treatment of strains with mutations in the primary targets can prove useful in identifying effects mediated by secondary drug targets, including the nature and extent of newly discovered and previously unsuspected pathways affected by the drug.

We describe here a method for drug target validation and the identification of secondary drug target effects that uses DNA microarrays to survey the effects of drugs on global gene expression patterns. We established that genetic and pharmacologic inhibition of gene function can result in extremely similar changes in gene expression. We also demonstrated that one can confirm a potential drug target by treating a deletion mutant defective in the gene encoding the putative target. Drug-mediated signatures from strains with mutations in pathways or processes directly or indirectly affected by the drug bore little or
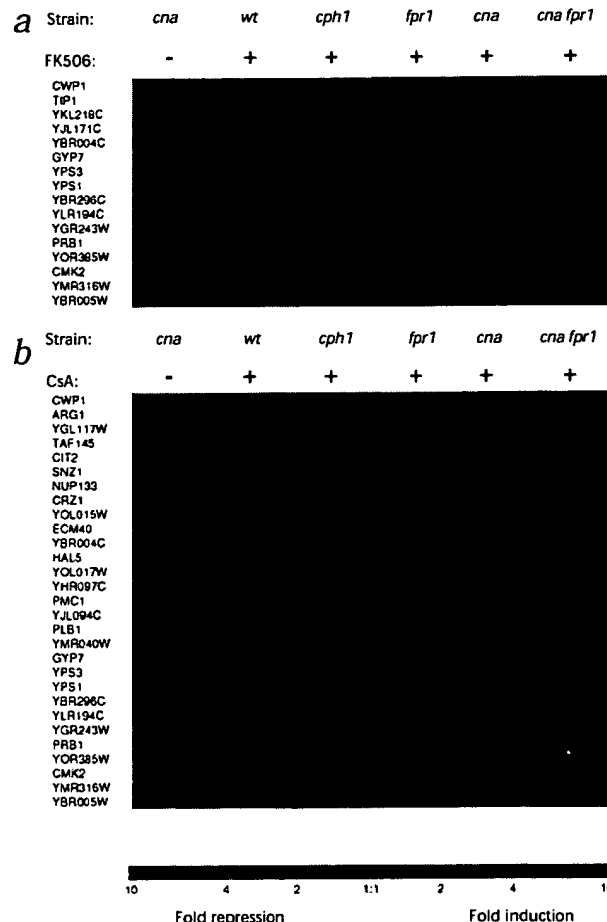
Fig. 5 Response of FK506 and CsA signature genes in strains with deletions in different genes. Genes with expression ratios greater than a factor of 1.8 in response to treatment with 1 µg/ml FK506 (*a*) or 50 µg/ml CsA (*b*) are listed (left side) and their expression ratios in the indicated strain are shown on the green (induction)–red (repression) color scale. *a*, Calcineurin (*cna*) mutant and FK506 treatment signature genes are in the first two columns. Almost all FK506 signature genes have expression ratios near unity in deletion strains involved in pathways affected by FK506 (calcineurin, *fpr1* and *cna fpr1* mutants) but not in deletion strains in unrelated pathways (*cph1*). *b*, Calcineurin (*cna*) mutant and CsA treatment signature genes are in the first two columns. Almost all CsA signature genes have expression ratios near unity in deletion strains involved in pathways affected by CsA (calcineurin, *cph1* and *cna cph1* mutants) but not in deletion strains in unrelated pathways (*fpr1*).

Strain:	wt	cna	fprl	cph1	cna fpr1	gen4
FK506:	+	+	+	+	+	+

CNA-dependent
YBR005W
YMR316W
CMK2

GCN4-dependent
MET18
CPA1
YOL119C
HOM2
YBR047W
SSU1
IDP1
ARG8
NIT1
YHM1
YOL164C
HAD1
CPA2
ADH5
YGL059W
YJR130C
ARO3
HIS3
PCL5
RIB5
SNO1
YHR029C
YBR147W
ARG3
HIS4
YGL117W
SNZ1
ARG5,6
ARO4
YPL250C
YER175C
ARG1

CNA- and GCN4-independent
URA1
HSC82
SSA2
OCV2
PDR5
SNQ2
HIS5

Complex behavior class
YER024W
YMR086W

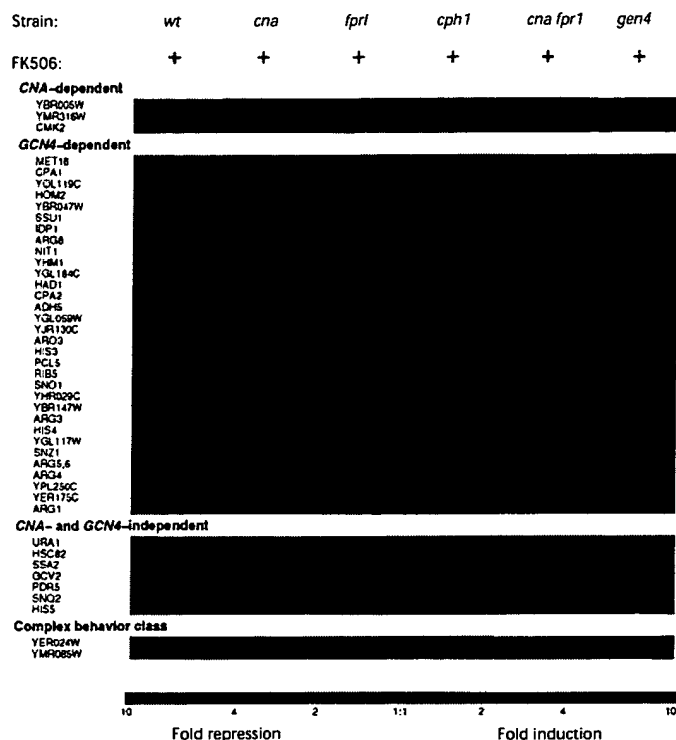Fold repression   Fold induction

Fig. 6 Response of FK506 signature genes in strains with deletions in different genes. Genes with expression ratios greater than a factor of 4 in at least one experiment are listed and their expression ratios in the indicated strain are shown in the green (induction)–red (repression) color scale. The genes have been divided into classes corresponding to these expected behaviors: 'CNA-dependent' genes respond to FK506 (50 μg/ml) except when either calcineurin genes or FPR1 or both are deleted; 'GCN4-dependent' genes respond to FK506 except when GCN4 is deleted. These genes still respond to FK506 when calcineurin genes or FPR1 or CPH1 are deleted; that is, their responses are not mediated by calcineurin, Cph1, or Fpr1. 'CNA- and GCN4-independent' genes respond to FK506 in all deletion strains tested. A 'complex behavior' class is provided for those genes that did not match the model of FK506 response mediated through calcineurin or Fpr1 or separately through Gcn4.

penile erection. It is possible that application of the 'decoder' to other compounds may show that they too have a potent activity against a target distinct from their intended target.

The ability to decode drug effects is dependent on the availability of functionally 'targetless' cells. In yeast, this is being achieved by systematically disrupting each yeast gene (Saccharomyces Deletion Consortium; http://sequence-www.stanford.edu/group/yeast_deletion_project/deletion.html). Efforts are underway to obtain expression profiles from each deletion mutant strain. Determining signatures resulting from inactivation of essential genes presents a unique problem, but it may be possible to do so by examining heterozygotes or by using a controllable promoter to reduce expression of the essential gene. Although it is already feasible to test several compounds in dozens of yeast strains, another challenge for the 'decoder' strategy will be the efficient selection of the mutants with deletions in genes most likely to encode the intended drug target. The signature correlation plots described are one metric that could be used as part of that selection process, but others need to be explored. Applying the 'decoder' to mammalian cells presents additional challenges. It is considerably more difficult to isolate functionally 'targetless' cells. Strategies involving titratable promoters, known specific inhibitors, anti-sense RNAs, ribozymes, and methods of targeting specific proteins for degradation are possible and should be tested. Another limitation is that not all cell types express the same set of genes and therefore 'off-target' effects may be different in different cell types. In addition, applying the 'decoder' to human cells will also require technical improvements that allow expression profiling from a small number of cells. Even the broader question of whether the insensitivity of 'off-target' signatures to the disruption of the main target is the exception or the rule can only be answered by the accumulation of more data. Barkai and Leibler, however, have argued in favor of robustness of biological networks, indicating that drug perturbations ('off-target' signatures) may be robust even when the system is subjected to another perturbation (such as a genetic disruption)(ref. 28). Many practical developments will be necessary if the 'decoder' concept is to be broadly applied.

Expression arrays have been used mainly as an initial screen for genes induced in a particular tissue or process of interest by focusing on genes with large expression ratios. We have found, however, that effort to refine experimental protocols and repeat experiments increases the reliability of the data and permits new applications. For example, it provides a larger set

no similarity to the wild-type drug expression profile. In contrast, drug-mediated signatures from strains with mutations in genes involved in pathways unrelated to the drug's action showed extensive similarity to the wild-type drug signature. By applying this approach to a drug that affects multiple pathways (FK506), we were able to decode a complex signature into component parts, including the identification of an 'off-target' signature that was mediated through pathways independent of calcineurin or the Fpr1 immunophilin.

Discussion

It is well-established that high-throughput biochemical screening can identify potent inhibitory compounds against a given target. The 'decoder' approach described here complements this process by evaluating the equally important property of specificity: the tendency of a compound to inhibit pathways other than that of its intended target. The ability to observe such 'off-target' effects will likely be useful in several ways. Profiling compounds with known toxicities will allow the development of a database of expression changes associated with particular toxicities. Recognition of potential toxicities in the 'off-target' signatures of otherwise promising compounds then may allow earlier identification of those likely to fail in clinical trials. Comparing the extent and peculiarities of 'off-target' signatures of promising drug candiates could provide a new way to group compounds by their effects on secondary pathways, even before those effects are understood. This may prove to be an alternative, potentially more effective, way to select compounds for animal and clinical trials. Some drugs are more effective against a related protein than against the originally intended target. Sildenafil (Viagra™), for example, was initially developed as a phosphodiesterase inhibitor to control cardiac contractility, but was found to be highly specific for phosphodiesterase 5, an isozyme whose inhibition overcomes defects in

Table 3 Yeast strains used

Strain	Relevant genotype	Reference
YPH499	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1	(34)
R563	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 his3::HIS3	(this study)
R558	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 fpr1::HIS3	(this study)
R567	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 cph1::HIS3	(this study)
MCY300	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 cna1Δ1::hisG cna2Δ1::HIS3	(21)
R132	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 cna1Δ1::hisG cna2Δ1::HIS3 cph1::kanr	(this study)
R133	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 cna1Δ1::hisG cna2Δ1::HIS3 fpr1::kanr	(this study)
R559	Mata ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 his3::HIS3 gcn4::LEU2	(this study)
BY4719	Mata trp1-Δ63 ura3-Δ0	(35)
BY4738	Matα trp1-Δ 63 ura3-Δ0	(35)
R491	Mata/α BY4719 X BY4738	(this study)
BY4728	Mata his3-Δ200 trp1-Δ63 ura3-Δ0	(35)
BY4729	Matα his3-Δ 200 trp1-Δ63 ura3-Δ0	(35)
R1226	Mata/α BY4728 X BY4729	(this study)

of genes at higher confidence levels that serve as a more unique signature for a given protein perturbation. In addition, it allows subtle signatures to be detected, when, for example, a protein is only partially inhibited. This may enable clinical monitoring of small changes in protein function in disease or toxicity states before they could otherwise be detected. Because the functions of many genes detected on transcript arrays are known, these microarrays are powerful tools that provide detailed information about a cell's physiology. For example, changes in the flux through a metabolic pathway are reflected in transcriptional changes in genes in the pathway[7]. Furthermore, it may be possible to indirectly measure protein activity levels from expression profiling data (S.F., *et al.*, unpublished data). Thus, although the eventual development of genomic methods allowing the direct measurement of all cellular protein levels will be an important achievement, transcript array technology offers an immediate and robust means of evaluating the effects of various treatments on gene expression and protein function.

Methods

Construction, growth and drug treatment of yeast strains. The strains used in this study (Table 3) were constructed by standard techniques[29]. To construct strain R559, strain R563 was transformed to Leu+ with plasmid pM12 digested by *Sal*I and *Mlu*I (provided by A. Hinnebusch and T. Dever). Strains R132 and R133 were constructed by transforming the bacterial kanamycin resistance cassette[30] flanked by genomic DNA from the *CPH1* and *FPR1* loci, respectively, and selecting for G418-resistant colonies. For experiments with FK506, cells were grown for three generations to a density of 1×10^7 cells/ml in YAPD medium (YPD plus 0.004% adenine) supplemented with 10 mM calcium chloride as described[31]. Where indicated, FK506 was added to a final concentration of 1 μg/ml 0.5 h after inoculation of the culture or to 50 μg/ml 1 h before cells were collected. CsA was used at a final concentration of 50 μg/ml. Cells were broken by standard procedures[32] with the following modifications: Cell pellets were resuspended in breaking buffer (0.2 M Tris HCl pH 7.6, 0.5 M NaCl, 10 mM EDTA, 1% SDS), vortexed for 2 min on a VWR multi-tube vortexer at setting 8 in the presence of 60% glass beads (425–600 μm mesh; Sigma) and phenol:chloroform (50:50, volume/volume). After separation of the phases, the aqueous phase was re-extracted and ethanol-precipitated. Poly A+ RNA was isolated by two sequential chromatographic purifications over oligo dT cellulose (New England Biolabs, Beverly, Massachusetts) using established protocols[32].

For experiments using 3-AT, wild-type or *his3/his3* cells were grown to early logarithmic phase in SC medium, pelleted and resuspended in SC medium lacking histidine for 1 hr in the presence or absence of 10 mM 3-

AT, as indicated. Cells were harvested and mRNA isolated as above. FK506 was obtained from the Swedish Hospital Pharmacy (Seattle, Washington) and purified to homogeneity by ethyl acetate extraction by J. Simon (Fred Hutchinson Cancer Research Center, Seattle, Washington). CsA was obtained from Alexis Biochemicals (San Diego, California); 3-AT was from Sigma.

Preparation and hybridization of the labeled sample. Fluorescently-labeled cDNA was prepared, purified and hybridized essentially as described[7]. Cy3- or Cy5-dUTP (Amersham) was incorporated into cDNA during reverse transcription (Superscript II; Life Technologies) and purified by concentrating to less than 10 μl using Microcon-30 microconcentrators (Amicon, Houston, Texas). Paired cDNAs were resuspended in 20–26 μl hybridization solution (3 × SSC, 0.75 μg/ml polyA DNA, 0.2% SDS) and applied to the microarray under a 22- × 30-mm coverslip for 6 h at 63 °C, all according to a published method[7].

Fabrication and scanning of microarrays. PCR products containing common 5' and 3' sequences (Research Genetics, Huntsville, Alabama) were used as templates with amino-modified forward primer and unmodified reverse primers to PCR amplify 6,065 ORFs from the *S. cerevisiae* genome. Our first-pass success rate was 94%. Amplification reactions that gave products of unexpected sizes were excluded from subsequent analysis. ORFs that could not be amplified from purchased templates were amplified from genomic DNA. DNA samples from 100-μl reactions were isopropanol-precipitated, resuspended in water, brought to a final concentration of 3× SSC in a total volume of 15 μl, and transferred to 384-well microtiter plates (Genetix Limited, Christchurch, Dorset, England). PCR products were spotted onto 1 × 3-inch polylysine-treated glass slides by a robot built essentially according to defined specifications[1,5,7] (http://cmgm.stanford.edu/pbrown/MGuide). After being printed, slides were processed according to published protocols[7].

Microarrays were imaged on a prototype multi-frame CCD camera in development at Applied Precision (Issaquah, Washington). Each CCD image frame was approximately 2-mm square. Exposure times of 2 s in the Cy5 channel (white light through Chroma 618–648 nm excitation filter, Chroma 657–727 nm emission filter) and 1 s in the Cy3 channel (Chroma 535–560 nm excitation filter, Chroma 570–620 nm emission filter) were done consecutively in each frame before moving to the next, spatially contiguous frame. Color isolation between the Cy3 and Cy5 channels was about 100:1 or better. Frames were 'knitted' together in software to make the complete images. The intensity of spots (about 100 μm) were quantified from the 10-μm pixels by frame-by-frame background subtraction and intensity averaging in each channel. Dynamic range of the resulting spot intensities was typically a ratio of 1,000 between the brightest spots and the background-subtracted additive error level. Normalization between the channels was accomplished by normalizing each channel to the mean intensities of all genes. This procedure is nearly equivalent to normalization between channels using the intensity

ARTICLES ·······

ratio of genomic DNA spots[7], but is possibly more robust, as it is based on the intensities of several thousand spots distributed over the array.

Signature correlation coefficients and their confidence limits. Correlation coefficients between the signature ORFs of various experiments were calculated using:

$$\rho = \sum_k x_k y_k \left/ \left(\sum_k x_k^2 \sum_k y_k^2\right)^{1/2}\right.$$

where x_k is the $\log_{10}$ of the expression ratio for the kth gene in the x signature, and y_k is the $\log_{10}$ of the expression ratio for the kth gene in the y signature. The summation is over those genes that were either up- or down-regulated in either experiment at the 95% confidence level. These genes each had a less than 5% chance of being actually unregulated (having expression ratios departing from unity due to measurement errors alone). This confidence level was assigned based on an error model which assigns a lognormal probability distribution to each gene's expression ratio with characteristic width based on the observed scatter in its repeated measurements (repeated arrays at the same nominal experimental conditions) and on the individual array hybridization quality. This latter dependence was derived from control experiments in which both Cy3 and Cy5 samples were derived from the same RNA sample. For large numbers of repeated measurements the error reduces to the observed scatter. For a single measurement the error is based on the array quality and the spot intensity.

Random measurement errors in the x and y signatures tend to bias the correlation towards zero. In most experiments, most genes are not significantly affected but do show small random measurement errors. Selecting only the '95% confidence' genes for the correlation calculation, rather than the entire genome, reduces this bias and makes the actual biological correlations more apparent.

Correlations between a profile and itself are unity by definition. Error limits on the correlation are 95% confidence limits based on the individual measurement error bars, and assuming uncorrelated errors[33]. They do not include the bias mentioned above; thus, a departure of ρ from unity does not necessarily mean that the underlying biological correlation is imperfect. However, a correlation of 0.7 ± 0.1, for example, is very significantly different from zero. Small (magnitude of ρ < 0.2) but formally significant correlation in the tables and text probably are due to small systematic biases in the Cy5/Cy3 ratios that violate the assumption of independent measurement errors used to generate the 95% confidence limits. Therefore, these small correlation values should be treated as not significant. A likely source of uncorrected systematic bias is the partially corrected scanner detector nonlinearity that differently affects the Cy3 and Cy5 detection channels.

The 1 μg/ml FK506 treatment signature was compared with more than 40 unrelated deletion mutant strain or drug signatures. These control profiles had correlation coefficients with the FK506 profile that were distributed around zero (mean ρ = –0.03) with a standard deviation of 0.16 (data not shown), and none had correlations greater than ρ = 0.38. Similarly, the calcineurin mutant strain signature correlated well with the CsA treatment signature (ρ = 0.71 ± 0.04) but not with the signatures from the negative controls (mean ρ = –0.02 with a standard deviation of 0.18).

Quality controls. End-to-end checks on expression ratio measurement accuracy were provided by analyzing the variance in repeated hybridizations using the same mRNA labeled with both Cy3 and Cy5, and also using Cy3 and Cy5 mRNA samples isolated from independent cultures of the same nominal strain and conditions. Biases undetected with this procedure, such as gene-specific biases presumably due to differential incorporation of Cy3- and Cy5-dUTP into cDNA, were minimized by doing hybridizations in fluor-reversed pairs, in which the Cy3/Cy5 labeling of the biological conditions was reversed in one experiment with respect to the other. The expression ratio for each gene is then the ratio of ratios between the two experiments in the pair. Other biases are removed by algorithmic numerical de-trending. The magnitude of these biases in the absence of de-trending and fluor reversal is typically about 30% in the ratio, but may be as high as twofold for some ORFs.

Expression ratios are based on mean intensities over each spot. Some smaller spots have fewer image pixels in the average. This does not degrade accuracy noticeably until the number of pixels falls below ten, in which case the spot is rejected from the data set. 'Wander' of spot positions with respect to the nominal grid is adaptively tracked in array subregions by the image processing software. Unequal spot 'wander' within a subregion greater than half-a-spot spacing is a difficulty for the automated quantitating algorithms; in this case, the spot is rejected from analysis based on human inspection of the 'wander'. Any spots partially overlapping are excluded from the data set. Less than 1% of spots typically are rejected for these reasons.

1. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470 (1995).
2. Schena, M. *et al.* Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. USA* 93, 10614–10619 (1996).
3. Shalon, D., Smith, S.J. & Brown, P.O. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* 6, 639–645 (1996).
4. Lockhart, D.J. *et al.* Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnol.* 14, 1675–1680 (1996).
5. DeRisi, J. *et al.* Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genet.* 14, 457–460 (1996).
6. Heller, R.A. *et al.* Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc. Natl. Acad. Sci. USA* 94, 2150–2155 (1997).
7. DeRisi, J.L., Iyer, V.R. & Brown, P.O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278, 680–686 (1997).
8. Lashkari, D.A. *et al.* Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl. Acad. Sci. USA* 94, 13057–13062 (1997).
9. Wodicka, L., Dong H., Mittman. M, Ho, M.-H. & Lockhart, D.J.. Genome-wide expression monitoring in Saccharomyces cerevisiae. *Nature Biotechnol.* 15, 1359–1367 (1997).
10. Cho, R.J. *et al.* A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell* 2, 65–73 (1998).
11. Gray, N.S. *et al.* Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* 281, 533–538 (1998).
12. Cardenas, M.E., Lorenz, M., Hemenway, C. & Heitman, J. Yeast as model T cells. *Perspect. Drug Discovery Design* 2, 103–126 (1994).
13. Klee, C.B., Ren, H. & Wang, X. Regulation of the calmodulin-stimulated protein phosphatase, calcineurin. *J. Biol. Chem.* 273, 13367–13370 (1998).
14. Tanida, I., Hasegawa, A., Iida, H., Ohya, Y. & Anraku, Y. Cooperation of calcineurin and vacuolar H(+)-ATPase in intracellular Ca2+ homeostasis of yeast cells. *J. Biol. Chem.* 270, 10113–10119 (1995).
15. Moser, M.J., Geiser, J.R. & Davis, T.N. Ca2+-calmodulin promotes survival of pheromone-induced growth arrest by activation of calcineurin and Ca2+-calmodulin-dependent protein kinase. *Mol. Cell. Biol.* 16, 4824–4831 (1996).
16. Mizunuma, M., Hirata, D., Miyahara. K., Tsuchiya, E. & Miyakawa, T. Role of calcineurin and Mpk1 in regulating the onset of mitosis in budding yeast. *Nature* 392, 303–306 (1998).
17. Yazdanbakhsh, K., Choi, J.W., Li, Y., Lau, L.F. & Choi, Y. Cyclosporin A blocks apoptosis by inhibiting the DNA binding activity of the transcription factor Nur77. *Proc. Natl. Acad. Sci. USA* 92, 437–441 (1995).
18. Molkentin, J.D. *et al.* A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell* 93, 215–228 (1998)
19. Mansuy, I.M., Mayford, M., Jacob, M., Kandel, E.R. & Bach, M.E. Restricted and regulated overexpression reveals calcineurin as a key component in the transition from short-term to long-term memory. *Cell* 92, 39–49 (1998).
20. Schreiber, S.L. & Crabtree, G.R. The mechanism of action of cyclosporin A and FK506. *Immunol. Today* 13, 136–142 (1992).
21. Cyert, M.S., Kunisawa, R., Kaim, D. & Thorner, J. Yeast has homologs (CNA1 and CNA2 gene products) of mammalian calcineurin, a calmodulin-regulated phosphoprotein phosphatase. *Proc. Natl. Acad. Sci. USA* 88, 7376–7380 (1991).
22. Jones, E.W. & Fink, G.R. in *The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression* (eds. Strathern, J.N., Jones, E.W. & Broach, J.R.) 181–299 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1982).
23. Hinnebusch, A. Translational regulation of yeast GCN4. *J. Biol. Chem.* 272, 21661–21664 (1997).
24. Hinnebusch, A.G. in *The Molecular and Cellular Biology of the Yeast*

Saccharomyces: Gene Expression. (eds. Jones, E.W., Pringle, J.R. & Broach, J.R.) 319–414 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1992).

25. Heitman, J. *et al.* The immunosuppressant FK506 inhibits amino acid import in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **13**, 5010–5019 (1993).
26. Balzi, E. & Goffeau, A. Yeast multidrug resistance: the PDR network. *J. Bioenerg. Biomembr.* **27**, 71–76 (1995).
27. Egner, R., Rosenthal, F.E., Kralli, A., Sanglard, D. & Kuchler, K. Genetic separation of FK506 susceptibility and drug transport in the yeast Pdr5 ATP-binding cassette multidrug resistance transporter *Mol. Biol. Cell* **9**, 523–543 (1998).
28. Barkal, N. & Leiber, S. Robustness in simple biochemical networks. *Nature* **387**, 913–917 (1997).
29. Schiestl, R.H., Manivasakam, P, Woods, R.A. & Gietz, R.D. Introducing DNA into yeast by transformation. *Methods: A companion to Methods in Enzymology* **5**, 79–85 (1993).
30. Wach, A., Brachat, A., Pohlmann, R. & Philippsen, P. New heterologous mod-

ules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. *Yeast* **10**, 1793–1808 (1994).
31. Garrett-Engele, P., Moilanen, B. & Cyert, M.S . Calcineurin, the Ca2+/calmodulin-dependent protein phosphatase, is essential in yeast mutants with cell integrity defects and in mutants that lack a functional vacuolar H(+)-ATPase. *Mol. Cell. Biol.* **15**, 4103–4114 (1995).
32. Ausubel, F.M. *et al.* in *Current Protocols in Molecular Biology* 13.12.1–13.12.5 (eds. Ausubel, F.M., *et al.*)(John Wiley & Sons, New York, 1993).
33. Bulmer, M.G. in *Principles of Statistics* 224–225 (Dover Publications, New York, 1979).
34. Sikorski, R.S. & Hieter, P. A system of shuttle vectors and yeast host strains designated for efficient manipulation of DNA in *Saccharomyces cerevisiae. Genetics* **122**, 19–27 (1989).
35. Brachmann, C.B. *et al.* Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115–132 (1998).